Mol Breeding (2010) 25:593–602
DOI 10.1007/s11032-009-9355-y

A SNP haplotype associated with a gene resistant to *Xanthomonas axonopodis* pv. *malvacearum* in upland cotton (*Gossypium hirsutum* L.)

Jinhua Xiao · David D. Fang · Muhammad Bhatti · Bill Hendrix · Roy Cantrell

Received: 21 September 2009 / Accepted: 23 November 2009 / Published online: 5 December 2009

Abstract An $F_{4:5}$ population of 285 families with each tracing back to a different F_2 plant, derived from a cotton bacterial blight resistant line 'DeltaOpal' and a susceptible line 'DP388', was artificially inoculated with bacterial blight *race 18* (*Xanthomonas axonopodis* pv. *malvacearum*) to assay their resistance or susceptibility to the disease. The segregation in the $F_{4:5}$ population indicates that the resistance was conditioned by a single dominant gene designated B_{12}. Simple sequence repeat (SSR) markers identified as putatively linked to the resistance gene by bulked segregant analysis were confirmed on the entire $F_{4:5}$ population. Three SSR markers, CIR246, BNL3545 and BNL3644 on chromosome 14, were found closely linked to B_{12}. The association between CIR246 and B_{12} was validated among 354 plants of 16 diverse varieties. Based on Monsanto SSR/single nucleotide polymorphism (SNP) consensus map, SNP markers closely linked to CIR246 were used to screen 'DeltaOpal' and 'DP388' for polymorphism. The polymorphic SNP markers were run on the $F_{4:5}$ population and the four SNP markers spanning 3.4 cM were found to flank the resistance gene on chromosome 14. The linkage between B_{12} and the 4-SNP marker haplotype was validated using 18 elite cotton lines. This 4-SNP marker haplotype can be used for marker assisted selection for bacterial blight resistance breeding programs or for screening germplasm collections for this locus rapidly.

Keywords Cotton (*Gossypium* spp.) · Bacterial blight · DNA markers · SSR markers · SNP marker haplotype · Marker-assisted selection · Bulked segregant analysis

Jinhua Xiao and David D. Fang contributed equally to this work.

J. Xiao (✉) · M. Bhatti · R. Cantrell
Monsanto Company, 800 N. Lindbergh Blvd, St. Louis, MO 63167, USA
e-mail: JINHUA.XIAO@STL.MONSANTO.COM

D. D. Fang
Cotton Fiber Bioscience Research Unit, USDA-ARS-SRRC, 1100 Robert E Lee Blvd, New Orleans, LA 70124, USA

B. Hendrix
Monsanto Company, Monsanto-Calgene Campus, 1920 East 5th St., Davis, CA 95616, USA

Introduction

Cotton (*Gossypium* spp.) is the leading natural fiber crop in the world. Bacterial blight, caused by *Xanthomonas axonopodis* pv. *malvacearum*, is a major disease of cotton occurring in most cotton-producing countries of the world, causing significant yield losses. In the US, yield losses caused by this disease in some years were as high as 3.4% (Brown 2001). In Australia, resistance to bacterial blight is a mandate for all commercial cotton varieties. The

pathogen enters the host plants through open stomata or wounds, and creates water-soaked lesions on leaves, stems, and bolls, followed by premature leaf senescence and reduced lint yield (Rungis et al. 2002). Disease management includes sanitary practices during ginning and seed processing, planting of acid-delinted and fungicide-treated seeds, destruction of residues from the previous crop, crop rotation, and use of resistant varieties. The deployment of resistant varieties is the most effective and economical means to control the disease and minimize yield loss.

Artificial inoculation is often required for the screening of resistant germplasm and selection of resistant individuals in breeding populations in order to develop resistant varieties. Natural infection can occur at a high level, but disease escapes are common among plants in an infected field. Traditional inoculation methods used for screening or phenotyping against bacterial blight can be time consuming, slow to respond, non-uniform in symptom development, and require controlled environments like a greenhouse or growth chamber. Due to the cost of such assays and the lack of reliable availability of inoculums, the utility of phenotype-based selection of plants to develop resistant varieties is limited in many breeding programs.

Cotton pathologists, geneticists and breeders have put tremendous efforts on the identification and characterization of different races, development of phenotyping assay methods, screening of germplasm for resistance sources, identification of different resistance genes, and the development of resistant varieties (Wright et al. 1998). So far, about 19 *Xcm* races have been identified in the USA (Verma 1986), and additional isolates have appeared in Africa. Of these races, race 18 is the most virulent and present in almost all cotton production areas in the world (Brown 2001). Resistance to bacterial blight has been studied extensively (Brinkerhoff 1970; Hillocks 1992). There are at least 22 reported resistance genes in cotton that confer differing degrees of resistance to various *Xcm* races carrying different avirulence genes in a typical gene-for-gene manner. Of these 22 genes, B_{12} confers a high level of resistance to all *Xcm* races presently found in the US and also others races found in Africa (Wallace and El-Zik 1989, 1990).

Molecular markers provide a novel avenue for the genotype-based resistance selection. Simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers have been proven very effective and provide powerful tools for mapping genes of interest for marker-assisted selection in breeding (Wang et al. 2006; Eathington et al. 2007). In cotton, several genes controlling disease resistance traits, including root-knot nematode [*Meloidogyne incognita* (Kofoid & White) Chitwood] (Shen et al. 2006; Wang et al. 2006; Ynturi et al. 2006), reniform nematode (*Rotylenchulus reniformis* Linford & Oliveira) (Dighe et al. 2009; Romano et al. 2009), verticillium wilt (*Verticillium dahliae* Kleb.) (Bolek et al. 2005), black root rot (*Thielaviopsis basicola*) (Niu et al. 2008), cotton leaf curl virus (Aslam et al. 2000) and cotton blue disease (Fang et al. 2009) have been tagged by molecular markers. As for tagging bacterial blight resistance genes, Wright et al. (1998) mapped *B2*, *B3*, *and b6* genes with restriction fragment length polymorphism (RFLP) markers. They also placed B_{12} gene on chromosome 14, but the closest linked RFLP marker was 11.4 cM. Later, Rungis et al. (2002) tried to tag B_{12} gene with amplified fragment length polymorphism (AFLP) and SSR markers, but failed to identify any closer markers. In this research, we report the identification and validation of SSR markers and SNP marker haplotype that are associated with the gene B_{12} conferring resistance to cotton bacterial blight *Race 18*.

Materials and methods

Mapping population and other plant materials

The resistant line 'DeltaOpal' (*G. hirsutum* L.), a commercial variety developed by Deltapine Australia Ltd, as a female, was crossed with the susceptible line 'DP388' in a greenhouse of Delta and Pine Land Company, Scott, Mississippi. An F_2 population of 285 plants derived from a single F_1 plant was advanced to F_4 through single seed decent. $F_{4.5}$ seeds were harvested from each one of the 285 F_4 plants.

Bacterial blight resistance assay of $F_{4.5}$ population

An isolate of *Xcm* (Race 18) was kindly provided by Dr. Peggy Thaxton at Texas A&M University. The isolate was sub-cultured one time on liquid LB medium and stored in a −80°C freezer. $F_{4:5}$ seeds were germinated in a paper towel for 24 h at 32°C.

The germinated seeds were transplanted into 4 × 10 cm cone-tubes filled with Metro-Mix 300 soil (SureGrow Horticulture, Bellevue, WA), and grown for 7 days in a greenhouse. Each F_5 family consisted of 21 seedlings. A total of about 7,000 seedlings were evaluated. According to Sedcole (1977), scoring 17 or more individuals will identify at least one susceptible plant in the progeny of a heterozygote in 99% of the test. Inoculation was made by scratching the back side of both cotyledons using a toothpick. The toothpick was dipped into a bacterial suspension ($\sim 10^6$ bacteria/ml) before each scratch (Thaxton and El-Zik 1993). After inoculation, seedlings were moved into a growth chamber set during the day at 28°C with lights and 20°C without lights at night. Humidity was set at 100% for day and night hours. The seedlings were arranged in a completely randomized design. Parental varieties were included in each experiment as controls. Bacterial blight symptom usually developed 4 days after inoculation. Disease symptoms were scored 7 days after inoculation. Resistant plants developed no necrosis in the area where inoculation occurred. Susceptible plants showed a water-soaked necrotic lesion and necrosis that spread beyond the site of inoculation (Fig. 1). Seedlings were scored as "resistant" (without water-soak symptom) or "susceptible" (with water-soak symptom). Scores were double checked by a second researcher.

After scoring, DNA was isolated from five plants from each $F_{4:5}$ family. A total of 1425 DNA samples were isolated.

Fig. 1 Bacterial blight evaluation (*Left*: susceptible, *Right*: resistant)

DNA isolation and bulked segregant analysis

Young healthy leaves were collected from plants of parental varieties, F_1, F_2 and $F_{4:5}$. For the F_1 and parental varieties, a bulk of 10 leaves from 10 plants was collected. For F_2 and F_5 seedlings, leaves were collected from individual plants. Total DNA was extracted from either fresh or frozen leaves using 2.0% hexadecyltrimethyl ammonium bromide according to Paterson et al. (1993). DNA was purified using Omega EZNA DNA isolation column (Omega Bio-Tek, Norcross, GA). To rapidly identify DNA markers associated with bacterial blight resistance, BSA was deployed as described by Michelmore et al. (1991). For the resistant bulk, DNA aliquots of 10 plants that were from 10 homozygous resistant F_5 families were pooled at equal ratio and diluted to 10 ng/μl. The susceptible bulk consisted of DNA aliquots from 10 plants that were from 10 homozygous recessive F_5 families. SSR primers that generated polymorphic patterns between bulks were retested using another two bulks, each composed of five F_5 individuals belonging to different F_5 families not included in the first two bulks. Those primers that still gave polymorphisms were further tested using the 30 individual DNA samples that were included in the bulks. The markers putatively linked to the resistance gene were analyzed in all 285 F_2 plants and 1425 individual $F_{4.5}$ seedlings.

SSR marker analysis

Two DNA bulks, 'DeltaOpal', 'DP388' and F_1 were analyzed with 4,247 pairs of SSR primers. Primer sequences for the public markers (BNL, CIR, JESPR, CM, and MGHES) can be obtained from Cotton Marker database (www.cottonmarker.org) (Blenda et al. 2006). The primer sequences of Monsanto markers (COT, C2, DC, DPL, and SHIN) are listed in Xiao et al. (2009). Forward primers were fluorescent-labeled with 6-FAM (6-carboxyfluorescein), HEX (4, 7, 2′, 4′, 5, 7-hexachloro-carboxyfluorescein) or NED (7′, 8′-benzo-5-fluoro 2′, 4, 7,-trichloro-5-carboxy-fluorescein). SSR primers were purchased from Sigma Genosys (Woodlands, Texas) or Applied Biosystems Inc. (Foster City, CA). Multiplex PCR was performed for all primers. Three pairs of primers with different dyes were multiplexed in each PCR reaction. The 10 μl PCR reaction included 20 ng

DNA, 2.5 μM each of the forward and reverse primers, 3.5 mM $MgCl_2$, 0.2 mM dNTPs, 1 unit of DNA Taq polymerase (Promega Corporation, Madison, WI), and 1× reaction buffer without $MgCl_2$. Amplification conditions were 95°C for 3 min, followed by 34 cycles of 94°C for 45 s, 55°C for 45 s, and 72°C for 1 min, with a final step of 72°C for 10 min. Amplified PCR products were separated and measured on an automated capillary electrophoresis system ABI 3730 XL (Applied Biosystems Inc.). GeneScan™-400 ROX™ (Applied Biosystems Inc.) was used as an internal DNA size standard. The output was analyzed with GeneMapper 3.7 software (Applied Biosystems Inc).

Development of SNP markers and SNP genotyping assay

All of the SNP markers used in this study were developed through a general genome-wide cotton SNP discovery method outlined below. The genomic DNA extracted from 'TM-1', a *G. hirsutum* line, was fully digested by *Pst I* or *Pvu II*. The restricted fragments were separated on agarose gel and fragments ranging from 0.7 to 3.0 kb were isolated from the gel and purified for the construction of reduced representative libraries. The libraries were sequenced and the clone sequences were checked for redundancy within and among the libraries. The primer pairs that amplify 500–800 bp of cotton genomic DNA were designed from the unique clone sequences. A panel of 22 diverse *G. hirsutum* lines was selected based on a previous SSR fingerprinting data (Xiao, unpublished data). These 22 lines are 'DeltaPearl', 'Sicala40' and 'Sicot189' from Australia; 'Jimian12' and 'Zhongmian35' from China; 'CIM473' and 'FH901' from Pakistan; 'Acala NemX', 'DP20B', 'DP5690', 'Explorer', 'FM989BGRR', 'HS26', 'NG3969R', 'Nucotn33B', 'PM2200RR', 'SG105', 'SG474', 'Sphinx', 'ST4892BR', 'ST5599BR' and 'TM-1'. Genomic DNA from these lines was amplified using the prime pairs and the resulting amplicons were sequenced. The amplicon sequences of these 22 lines amplified from the same primer pair were compared through sequence alignment to identify putative SNPs (Larkin et al. 2007). The minor allele had to have at least two counts among the 22 lines in order to consider it as a putative SNP. Once a SNP was identified, an end-point TaqMan® assay was developed to discriminate SNP alleles by properly designing primers and probes using design tools offered by Applied Biosystems Inc. (https://www2.appliedbiosystems.com/support/software/assaysbydesign). The 7 μl TaqMan® assay reaction contained 10 μM each of primers, 0.2 μM each of probes, 5 ng genomic DNA, and 1× TaqMan® universal PCR Master Mix (Applied Biosystems Inc.). PCR temperature profiles were: 50°C 2 min, 95°C 10 min, followed by 40 cycles of 92°C 15 s and 60°C 1 min. At the end of PCR, plates were scanned using ABI 7300 Real-Time PCR System, and SNP alleles plotted using software SDS 1.1 (Applied Biosystems Inc.). All SNP markers were validated for segregation in a Monsanto interspecific F_2 (discovery) population with 94 individual plants derived from 'DP33B' (*G. hirsutum*) and 'GB679' (*G. barbadense*) (Xiao et al. 2009) with public anchor SSR markers. SNPs segregating in a normal 1:2:1 manner were validated, since others potentially could be sequence polymorphisms arising between the A-genome and D-genome (pseudo-SNPs). The SNP attributes are presented in Table 3. All SNP sequences found linked to the trait of interest were deposited in GenBank.

Selection of diverse germplasm for SSR markers, SNP marker haplotype validation and phenotype confirmation

After the identification of SSR markers closely linked to the resistance gene, more than 1,000 cotton germplasm lines were screened using the SSR markers and classified as "resistant", "segregate", or "susceptible" purely based on the marker genotypes. The marker validation process was done in two steps. The first step was to validate the SSR markers. The second step was to validate the SNP marker haplotype.

For the first step, 16 varieties (Table 2) were assayed for their reactions to *Xcm Race 18* infection. The bacterial blight resistance evaluation was performed as described above.

For the second step, a set of nine lines having the 'resistance' SSR alleles and nine lines having the 'susceptible' SSR alleles were selected for genotyping at the four SNP markers and resistant or susceptible phenotype confirmation. Two methods were used to evaluate bacterial blight resistance

evaluation. The first method was by dipping a toothpick in a solution of Xcm (10^6 bacteria/ml) and creating a scratch on the underneath surface of a true leaf (not a cotyledon). The second method was to use an Agro-Jet inoculating gun at 50 PSI to inject 0.5 ml (10^6 bacteria/ml) of a bacterial solution on the underside of the leaf. A plexiglass solid surface was used to support the leaf in place against the pressure of the inoculating gun. The gun was adjusted to deliver 0.5 ml of solution per trigger. Ten plants of each of the 18 lines were inoculated at 4th leaf stage at two locations by the gun method and one location using scratch inoculation method. The inoculated plants were kept in a growth chamber at 28°C with a 16 h light cycle at a 30% relative humidity.

Linkage analysis

Segregation data for bacterial blight resistance and SSR markers from 285 F_2, and F_5 families were mapped using program JoinMap3.0 (Van Ooijen and Voorrips 2001) with LOD score $\geq$ 5.0. Chi-square tests were used to check segregation of markers and disease resistance gene against an expected 1:2:1 frequency.

Results

Inheritance of bacterial blight resistance in 'DeltaOpal'

In the course of this research, more than 100 seedlings of 'DeltaOpal' were evaluated, all showed resistance to *Xcm Race 18* without any exception. F_1 seedlings were also resistant. 'DP388' was susceptible. Of the 285 $F_{4.5}$ families, 117 showed resistances to the disease (RR), 122 susceptible to the disease (rr), and 46 segregated for the resistance within the family (Rr). The segregation fits into 7:2:7 (RR:Rr:rr) ratio with a χ^2 value of 3.55 ($P > 0.1000$). For any given heterozygous (Rr) F_5 family, the segregation usually fits 3 (resistant):1 (susceptible). These results clearly indicated that the bacterial blight resistance in 'DeltaOpal' was controlled by a single dominant gene. Because this gene confers resistance to *Xcm Race 18* in dominant manner, it should be the gene B_{12} (Wright et al. (1998); Rungis et al. 2002). Our mapping results would support this (see below).

Identification of SSR markers linked to B_{12}

Of the 4,247 SSR markers screened, 265 (6.24%) were polymorphic between 'DeltaOpal' and 'DP388'. Twenty-six markers revealed difference between two DNA bulks. After analyzing them in the second pair of bulks and 30 individuals comprising the bulks, only 4 markers (CIR246, BNL3545, BNL3644, and BNL1403) were found putatively linked to the resistance locus. Then all 1425 F_5 plants with known bacterial blight phenotypes and 285 F2 plants were analyzed with these markers. The linkage map of the B_{12} locus region is shown in Fig. 2.

Except marker CIR246 that segregated at one single locus, the other three markers revealed two or more loci. The marker fragments in 'DeltaOpal' and 'DP388' are listed in Table 1. CIR246 was previously mapped only on chromosome 14 by Nguyen et al. (2004), Guo et al. (2007), and He et al. (2007). Recently, Xiao et al. (2009) mapped CIR246, BNL3644 and BNL3545 on chromosome 14 as well. Thus, we conclude that the bacterial blight resistance gene B_{12} in 'DeltaOpal' resides on chromosome 14. Our result is consistent with that of Wright et al. (1998) who located the B_{12} gene on chromosome 14 which was supported by Rungis et al. (2002).

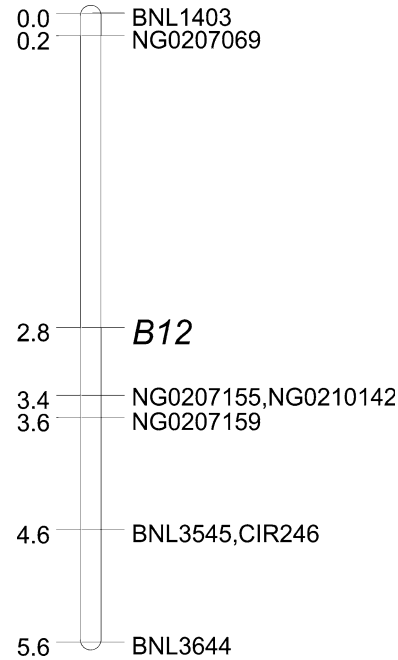

Fig. 2 Genetic map of B_{12} region on chromosome 14 (D02)

Table 1 SSR markers linked to B_{12} and the fragments in 'DeltaOpal' & 'DP388'

Markers	Repeats	DeltaOpal[a]	DP388[a]
BNL1403	AG	167	167
		170	170
		179	
			183
		187	187
BNL3545	CA	115	115
		119	
			129
		183	183
CIR246	CA	146	
			166
BNL3644	AG	184	184
			190
		194	

[a] DNA fragment size in bp

Validation of the association between SSR markers CIR246 and B_{12} in 16 varieties

After the SSR markers closely linked to the B_{12} were identified, we screened more than 1,000 cotton varieties using markers CIR246 (data not shown). We only screened marker CIR246 because of: (1) closest marker, and (2) one single locus. Sixteen varieties were selected, and their reaction to *Xcm Race 18* infection was evaluated. DNA was isolated from each individual plant, and CIR246 marker genotype was compared with the phenotypic score. The results are shown in Table 2. At the marker CIR246 locus, besides allele 166 bp present in 'DP388', a new allele 156 bp was observed in 9 varieties. This allele is linked to the susceptible allele at B_{12} locus. In cotton breeding, it is a common practice to bulk sibling progeny lines (usually at F_4 or $F_{4:5}$ stage) with similar agronomic performance before forming a variety. Consequently, a cotton variety may not be a pure line, rather a mixture of multiple sibling lines. As can be seen from Table 2, eight varieties contained two alleles at CIR246 marker locus. However, due to inbreeding, all individual plants within any given variety should be homozygous at a particular locus. We analyzed all 354 plants with CIR246, and did not observe a single plant that was heterozygous. The marker results completely matched the disease resistance evaluation scores among all 354 plants of 16 varieties. This suggested that the marker CIR246 was very closely linked to the B_{12} gene across a wide array of genetic backgrounds.

Table 2 Sixteen cotton genotypes used to validate the association between CIR246 and B_{12}

Order	Genotype	Alleles at CIR246 locus	B_{12} locus prediction based on marker	Inoculation results	
				Resistant	Susceptible
1	03Q060	146	RR	21	0
2	DPX 3163	146	RR	15	0
3	01V25	146&156	RR & rr	5	15
4	PMX 1144	146&156	RR & rr	11	5
5	00H26	146&166	RR & rr	2	15
6	03V42	146&166	RR & rr	8	9
7	DP660	146&166	RR & rr	75	48
8	00Z10	156	rr	0	20
9	01Z34	156	rr	0	17
10	389006-9105-103-201	156	rr	0	10
11	Maxxa	156	rr	0	12
12	01X02	166	rr	0	12
13	03Z012	166	rr	0	21
14	389003-9103-103-203	156 &166	rr	0	12
15	DPX9269	156&166	rr	0	16
16	Eva	156&166	rr	0	5

Identification of SNP markers linked to B_{12}

Based on Monsanto SSR/SNP consensus map that has about 7,000 markers, SNP markers within 10 cM from CIR246 and BNL3545 on either side were selected to screen the two parents for polymorphism that then were analyzed on the entire F_5 population. Four SNP markers, NG0207159, NG0207155, NG0210142, and NG0207069 spanning 3.4 cM on chromosome 14, were found to flank the resistance gene (Fig. 2). The attributes of these four SNP markers are shown in Table 3. The F_5 families having the resistance haplotype, A-C-T-T, are all resistant to the disease, and the F_5 families having the susceptible haplotype, G-G-C-A, are all susceptible to the disease.

Validation of the linkage between the SNP haplotype and B_{12} in 18 elite cotton lines

Nine lines, '01Q12', '03Q066', '08P652', '01Q08', '00Q05', '08Q160', '03Q075', '00X16B' and '03W306R' that have the "resistance" SSR allele (146 bp at CIR246 locus), and nine lines, '03X179R', '01Z20', '01X06', '01X04', '03X164R', '04Z007', '04Z045', '01X02' and '01Z18' that possess the "susceptible" SSR alleles (156 or 166 bp at CIR246 locus), were selected for genotyping at the four SNP markers and resistant or susceptible phenotype confirmation in a growth chamber. All the nine lines having the "resistance" SSR allele have the same SNP haplotype, A-C-T-T, that is identical to the one in 'DeltaOpal', and were resistant to the disease confirmed by the artificial inoculation assay. All the nine lines possessing the "susceptible" SSR alleles have the same SNP haplotype, G-G-C-A, that is identical to the one in 'DP388', and were susceptible to the disease confirmed by the artificial inoculation assay.

Discussion

Upland cotton is an allotetraploid with 26 pairs of chromosomes. It has two sub-genomes, *At* (chromosomes 1–13) and *Dt* (chromosomes 14–26). Due to sequence homology between these two sub-genomes, many SSR primers revealed duplicate loci present in both sub-genomes (Guo et al. 2007; Nguyen et al.

Table 3 Attributes of SNP markers associated with haplotype for cotton bacterial blight resistance

Marker	NG0207069	NG0210142	NG0207155	NG0207159
Position (cM)	0.2	3.4	3.4	3.6
Resistant allele	A	C	T	T
Susceptible allele	G	G	C	A
Fwd_primer	CCCTCTCCCTCTA	GGTAGGTTTTCTGTT	CCAAAGTTGAGAGCATT	GGTCAGTGATAGGAGT
	CCCTTGATAAAG	GGCTTTTCAT	TCGTTGAA	TCAAAAGGT
Rev_primer	CCAAGCATTCAACTT	GCAGGATGGGAGAGGGCTA	GCCCAAGTGGTAGCA	CGGTTTCTCAAGCTATACT
	AGTGACCTATAGA		TAATTGTC	GATCATCA
Vic_probe	ATCCCAATCATCAAGC	TCCTTGAACCCCACAATA	CTTGGTAGGTGATCGTGT	TTTGGATTCGATATAAAAT
Fam_probe	ATCCCAATCGTCAAGC	CTTGAACGCCACAATA	TTGGTAGGTGATCATGT	TTTGGATTCGATTTAAAAT

2004; Xiao et al. 2009). When using molecular markers to assist breeding, it is ideal that the marker associated with the trait segregates only at one locus. This is the case for the SSR marker CIR246 that is specific to *Dt* sub-genome chromosome 14. However, the other three SSR markers (BNL1403, BNL3545, and BNL3644) revealed two loci. For example, besides one locus at chromosome 14, the SSR marker BNL3545 also has one locus in chromosomes 2 (Xiao et al. 2009) that is homologous to chromosome 14. Because the BNL3545 marker locus at chromosome 14 completely co-segregated with the CIR246 locus after analyzing at least 5,000 DNA samples, we suspected that these two might be the same locus. We cloned BNL3545 fragments 115, 119, 129 and 183 bp, and CIR246 fragments 146, 156 and 166 bp from 15 diverse cotton varieties. About 250 clones were sequenced (data not shown). After aligning all sequences, it was found that BNL3545 and CIR246 belonged to the same contig, and the primer sequences were about 100 bp apart. Thus, we conclude that the BNL3545_119 bp/129b locus is the same as the CIR246_146 bp/166 bp locus. Due to different priming sites, BNL3545 amplified an additional locus at chromosome 2, while CIR246 only revealed one locus.

Resistant germplasm that are elite and locally adapted are the most desirable as the resistant parents. Identification of such resistance germplasm is very important in the breeding for disease resistance. Screening a large number of accessions through traditional artificial inoculation to identify resistant ones is very costly and time-consuming. Genotyping the elite lines for the markers flanking the resistance gene can identify the putative resistant lines based on their haplotypes at the resistance gene region. We conducted the phenotype validation of 18 elite lines, nine lines having the resistance haplotype and nine lines possessing susceptible haplotype, and found the phenotype of these 18 lines matched their corresponding haplotype without any exception. This suggests that the resistance haplotype defined by these four SNP markers can be used for the identification of resistant lines or accessions in a wide range of *G. hirsutum* germplasm through marker-assisted screening.

Traditional artificial inoculation methods to screen individuals resistant to bacterial blight in the breeding populations require controlled environments like greenhouses and growth chambers and are very time-consuming and of high cost, limiting the application of such phenotype-based resistance selection to only small number of breeding populations or only some advanced breeding lines. In breeding of resistance that is controlled by a single gene like the B_{12} gene for bacterial blight resistance, selection of resistant individuals is most desirable in early generations like F_2 or BC_1F_1 if possible. Selection at such an early stage when the population is still segregating is only possible by the application of markers tightly linked to the resistant gene. A SNP haplotype that flanks the resistant gene greatly facilitates the selection of the resistance genotypes. Marker-assisted selection for resistant individuals can be performed at any development stage of the plant including seeds, and even has a low cost and fast turnaround of information compared to the phenotype-based selection.

One of biggest challenges in breeding for resistant varieties is that pathogens can mutate. Mutations that allow the pathogen to reproduce on a variety that is resistant to all other races of that pathogen can quickly become the dominant race and that variety becomes susceptible. The wide deployment of a single source of resistance has historically resulted in crops becoming vulnerable to epidemics caused by a mutated pathogen. Therefore, identification and utilization of novel resistance sources is the major strategy to prevent the epidemic caused by newly mutated pathogens. The resistance among the different lines or accessions can be conditioned by the same resistance allele, different resistance alleles of the resistance locus, or resistance alleles at different loci. Genotyping resistant lines confirmed by phenotype at the markers defining the resistance haplotype can lead to the identification of novel resistance alleles or locus/loci. Putative novel resistance locus can be confirmed by allelism test in segregating populations derived from a resistant parent with known resistance and a candidate with a novel locus.

Cotton is not a restricted self-pollinated crop species and has a cross-pollination rate up to 50% depending on insect population in an open field. Varieties are not true inbreds, but heterogeneous in most cases. Homogeneity at the resistance locus of resistance varieties to be deployed in the areas where the disease is the major threat to the cotton production is very important. The SNP haplotype for cotton

bacterial blight resistance reported here can be used to monitor the seed purity for the resistance to the disease.

The haplotype for cotton bacterial blight resistance reported here spans a distance of 3.4 cM. That is short enough for marker-assisted screening of resistant lines or accessions from germplasm collections, marker-assisted selection of resistant individuals from breeding populations, identification of novel resistance alleles at the resistance locus and novel resistance loci if existing, and monitoring seed purity for the resistance of resistant varieties. The haplotype is also a start point for further defining a finer location of the resistance gene that may result in the eventual cloning of the resistance gene by map-based and candidate gene approaches.

Acknowledgments We thank Dr. Terry Wheeler at Texas A&M University and Dr. John Yu at USDA for their review comments for improving the manuscript.

Disclaimer Monsanto Company assumes no liability or claims associated with any use of the sequences or DNA markers disclosed in this paper. Users of the sequences or DNA markers may acknowledge Monsanto Company as the source of the information. Mention of a trademark, proprietary product or vendor does not constitute a guarantee by the US Department of Agriculture, and does not imply approval or recommendation of the product to the exclusion of others that may be suitable.

References

Aslam M, Jiang C, Wright R, Paterson AH (2000) Identification of molecular markers linked to leaf curl virus disease resistance in cotton. J Sci I R Iran 11:277–280

Blenda A, Scheffler J, Scheffler B, Palmer M, Lacape J-M, Yu J, Jesudurai C, Jung S, Muthukumar S, Yellambalase P, Ficklin S, Staton M, Eshelman R, Ulloa M, Saha S, Burr B, Liu S, Zhang T, Fang D, Pepper A, Kumpatla S, Jacobs J, Tomkins J, Cantrell R, Main D (2006) CMD: a cotton microsatellite database resource for *Gossypium* genomics. BMC Genomics 7:132

Bolek Y, El-Zik KM, Pepper AE, Bell AA, Magill CW, Thaxton PM, Reddy OUK (2005) Mapping of verticillium wilt resistance genes in cotton. Plant Sci 168:1581–1590

Brinkerhoff LA (1970) Variation in *Xanthomonas malvacerum* and its relation to control. Annu Rev Phytopathol 8:85–110

Brown JK (2001) Bacterial blight. In: Kirkpatrick TL, Rothrock CS (eds) Compedium of cotton diseases, 2nd edn. APS Press, St. Paul, pp 34–35

Dighe ND, Robinson AF, Bell AA, Menz MA, Cantrell RG, Stelly DM (2009) Linkage mapping of resistance to reniform nematode in cotton following introgression from *Gossypium longicalyx* (Hutch. & Lee). Crop Sci 49:1151–1164

Eathington SR, Crosbie TM, Edwards MD, Reiter RS, Bull JK (2007) Molecular markers in a commercial breeding program. Crop Sci 47:S154–S163

Fang DD, Xiao J, Canci PC, Cantrell RG (2009) A new SNP haplotype associated with blue disease resistance gene in cotton (*Gossypium hirsutum* L.). Theor Appl Genet (in press)

Guo WZ, Cai CP, Wang CB, Han ZG, Song XL, Wang K, Niu XW, Wang C, Lu KY, Shi B, Zhang TZ (2007) A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in *Gossypium*. Genetics 176:527–541

He DH, Lin ZX, Zhang XL, Nie YC, Guo XP, Zhang YX, Li W (2007) QTL mapping for economic traits based on a dense genetic map of cotton with PCR- based markers using the interspecific cross of *Gossypium hirsutum* x *Gossypium barbadense*. Euphytica 153:181–197

Hillocks RJ (1992) Bacterial blight. In: Hillocks RJ (ed) Cotton disease. CAB International, Wallingford, UK, pp 39–85

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) ClustalW and ClustalX version 2. Bioinformatics 23:2947–2948

Michelmore RW, Paran I, Kessell RV (1991) Identification of markers linked to disease-resistance genes bybulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. PNAS 88:9828–9832

Nguyen TB, Giband M, Brottier P, Risterucci AM, Lacape JM (2004) Wide coverage of the tetraploid cotton genome using newly developed microsatellite markers. Theor Appl Genet 109:167–175

Niu C, Lister HE, Nguyen B, Wheeler TA, Wright RJ (2008) Resistance to *Thielaviopsis basicola* in the cultivated a genome cotton. Theor Appl Genet 117:1313–1323

Paterson AH, Brubaker CL, Wendel JF (1993) A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. Plant Mol Biol Report 11:122–127

Romano GB, Sacks EJ, Stetina RS, Robinson AF, Fang DD, Gutierrez OA, Scheffler JA (2009) Identification and genomic location of a reniform nematode (*Rotylenchulus reniformis*) resistance locus (Ren^{ari}) introgressed from *Gossypium aridum* into upland cotton (*G. hirsutum*). Theor Appl Genet 120:139–150

Rungis D, Llewellyn D, Dennis ES, Lyon BR (2002) Investigation of the chromosomal location of the bacterial blight resistance gene present in an Australian cotton (*Gossypium hirsutum* L.) cultivar. Aust J Agri Res 53:551–560

Sedcole JR (1977) Number of plants necessary to recover a trait. Crop Sci 17:667–668

Shen X, Van Becelaere G, Kumar P, Davis RF, May LO, Chee P (2006) QTL mapping for resistance to root-knot nematodes in the M-120 RNR Upland cotton line (*Gossypium*

hirsutum L.) of the Auburn 623 RNR source. Theor Appl Genet 113:1539–1549

Thaxton PM, El-Zik KM (1993) Methods for screening and identifying resistance to the bacterial blight pathogen in cotton in the MAR program. Proc Beltwide Cotton Prod Res Conf Cotton Dis Counc 53:211–212

Van Ooijen JW, Voorrips RE (2001) JoinMap® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Verma JP (1986) Bacterial blight of cotton. CRC Press Inc, Boca Raton, FL

Wallace TP, El-Zik KM (1989) Inheritance of resistance in three cotton cultivars to the HVI isolate of bacterial blight. Crop Sci 29:1114–1119

Wallace TP, El-Zik KM (1990) Quantitative analysis of resistance in cotton to three new isolates of the bacterial blight pathogen. Theor Appl Genet 79:443–448

Wang C, Ulloa M, Roberts PA (2006) Identification and mapping of microsatellite markers linked to a root-knot nematode resistance gene (*rkn1*) in Acala NemX cotton (*Gossypium hirsutum* L.). Theor Appl Genet 112:770–777

Wright RJ, Thaxton PM, El-Zik KM, Paterson AH (1998) D-subgenome bias of Xcm resistance genes in tetraploid *Gossypium* (cotton) suggests that polyploidy formation has created novel avenues for evolution. Genetics 149:1987–1996

Xiao J, Wu K, Fang DD, Stelly DM, Yu J, Cantrell RG (2009) New SSR markers for use in cotton (*Gossypium* spp.) improvement. J Cotton Sci 13:75–157

Ynturi P, Jenkins JN, McCarty JC Jr, Gutierrez OA, Saha S (2006) Association of root-knot nematode resistance genes with simple sequence repeat markers on two chromosomes in cotton. Crop Sci 46:2670–2674